What is a Geographic Information System (GIS)?

Basic concepts and applications of GIS for criminal justice and policing

The 7th Annual International Crime Mapping Research Conference

Wednesday, March 31 (2:30 – 4 pm)

(*1 ½ hour workshop*)

Katie M. Filbert

Workshop Objectives

What is a GIS?
- Definition
- Components

Fundamental Mapping & GIS Concepts
- Map scale
- Map projections
- Symbolization
- GIS Data Models

Thematic Mapping
- Proportional Symbols
- Choropleth Maps

GIS Analysis
- Queries
    - Spatial queries
    - Attribute/Tabular queries
- Buffers/Proximity
- Spatial Overlay
    - Point-in-polygon
    - Line-on-polygon
    - Polygon overlay
- Spatial patterns
    - Clustered/Random/Regular (evenly spaced)
- Spatial diffusion/displacement

GIS Applications for Law Enforcement
- Operations
    - Resource/staffing management
    - Vehicle tracking/routing
- Crime analysis
    - Problem-oriented policing
- Citizen awareness
    - Crime statistics
    - Sex offender registry

What is a GIS?

In essence, GIS has three major capabilities (computer mapping, spatial analysis and spatial database). The software systems run in context of the computer hardware (for processing, manipulating and displaying information) and communications/data networks (for transmitting information). In broader context, GIS involves people and their institutions (police department). The stakeholders not only include the GIS/crime analysts, but also the developers who put together the GIS system, the managers that oversee the project and how GIS fits into the organization, and the users (officers). The results from GIS analysis can inform both policy and decision makers and the general public, the constituents.

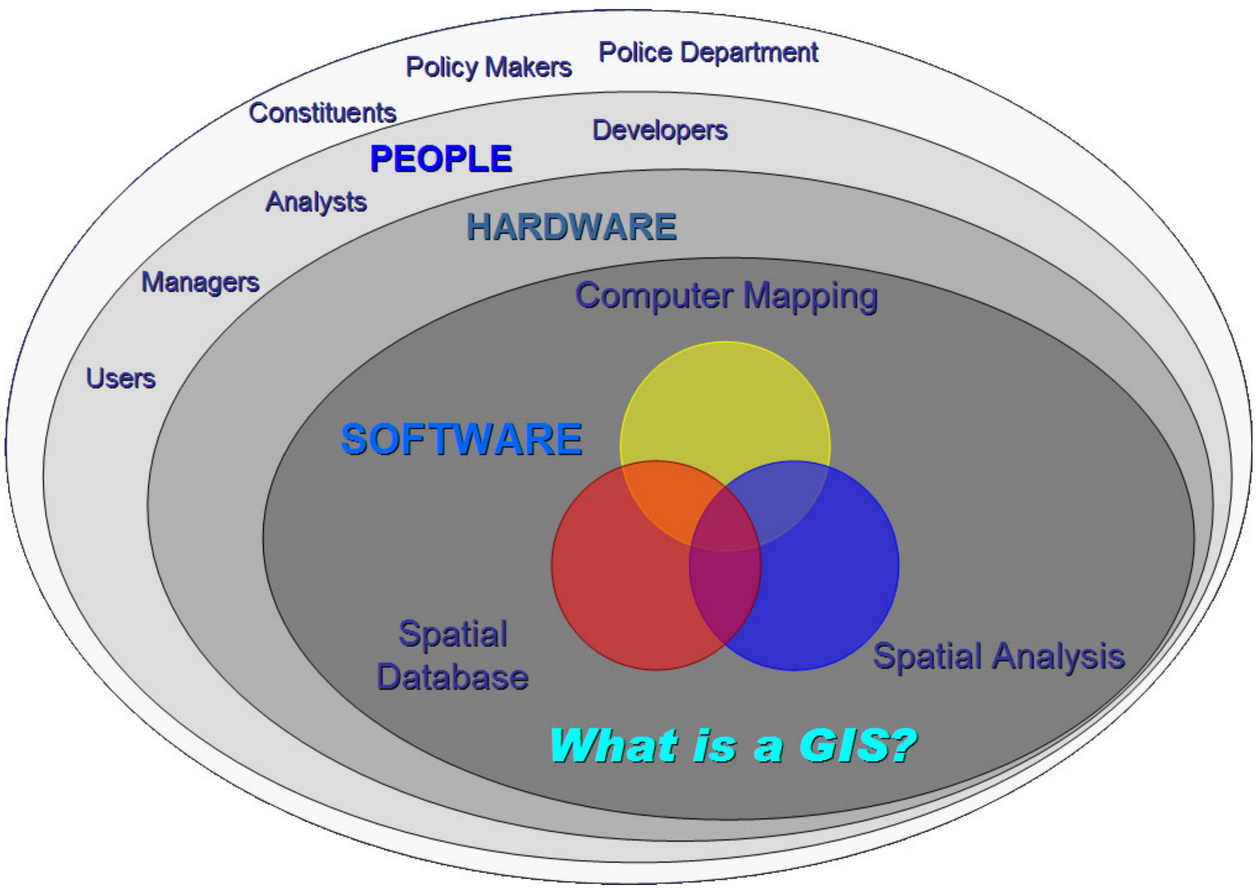


Figure 1: What is a GIS?

GIS Data Sources

(far from comprehensive list)

- Local governments
    - City of Boston:
        - http://www.cityofboston.gov/bra/maps/maps.asp
    - City of Cambridge:
        - http://www.cambridgema.gov/GIS/

- State government
    - Massachusetts GIS clearinghouse
        - http://www.state.ma.us/mgis/

- Federal government
    - Geodata.gov (GIS clearinghouse)
    - US Census Bureau
        - http://www.census.gov/geo/www/tiger/index.html
    - USGS (United States Geological Survey)
        - http://www.usgs.gov/pubprod/
        - http://edcsns17.cr.usgs.gov/EarthExplorer/
    - Other
        - FEMA
            - http://www.gismaps.fema.gov/
        - EPA
            - http://www.epa.gov/epahome/gis.htm
    - Statistics
        - FedStats
            - http://www.fedstats.gov/
        - CDC (Center for Disease Control) – health statistics
            - http://www.cdc.gov
        - Bureau of Justice Statistics
            - http://www.ojp.usdoj.gov/bjs/
        - Bureau of Labor Statistics
            - http://www.bls.gov/gps/

- Universities
    - University of Michigan – statistics
        - http://www.icpsr.umich.edu/access/index.html

- Commercial sources
    - Ikonos
        - http://www.ikonos.com
    - ESRI
        - http://www.esri.com/data/
    - NavTech
        - http://www.navteq.com/
- Custom data sources

Mapping & GIS Fundamentals

Map scale

Maps are scale models of the 'real world'. A map usually features a scale bar, used to measure distances on the map. It may also provide a ratio scale, such as 1:100,000, which tells you the map is reduced 100,000 times from reality; Thus, one inch on the map equals 100,000 inches in reality. 100,000 inches equates to 8,333 ft or roughly 1 ½ miles.

As scale models of the world, we cannot possibly depict all the details and complexities of the real world. We must pick and choose which features to include in the GIS while omitting many lesser details. The 1:10,000 map shows all the streets while the 1:35,000 only shows major streets. The 1:35,000 shows important buildings such as the hospital and train station, while the 1:100,000 map must omit those. From the 1:100,000 map to the 1:400,000 map, Bow River goes from an area (polygon) feature to a linear (line) feature, and the river loses some of its intricate details through **generalization**. Finally, from 1:400,000 to 1:600,000, major attractions such as Cave & Basin and Sulphur Mountain Gondola are omitted.

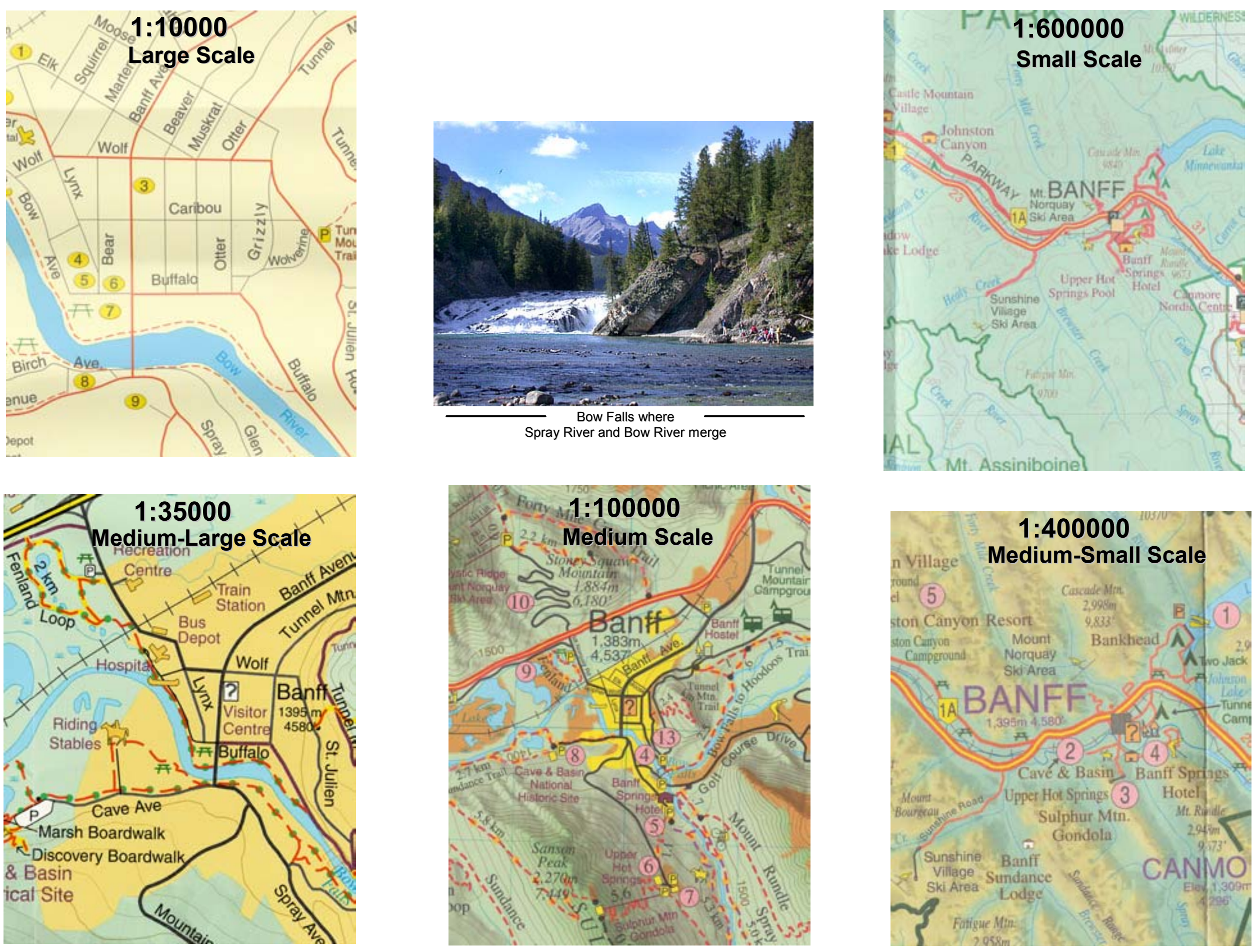


Figure 2: Banff, Alberta (Canadian Rockies) at different map scales

Map Projections

The Jasper map also indicates a map projection, Universe Transverse Mercator (UTM) with parameters, Zone 11 and Datum NAD83. To understand how UTM coordinates translate to actual spatial locations, let's consider map projections for world maps.

Map projections are used to 'flatten' the 3-dimensional world onto a 2d surface such as paper or the computer screen. Some distortion on the map is inevitable. We can select from numerous map projections to find one suited for a particular area, for a map or GIS data. Some projections are suited for world maps while others for small portions of a U.S. state or other area.

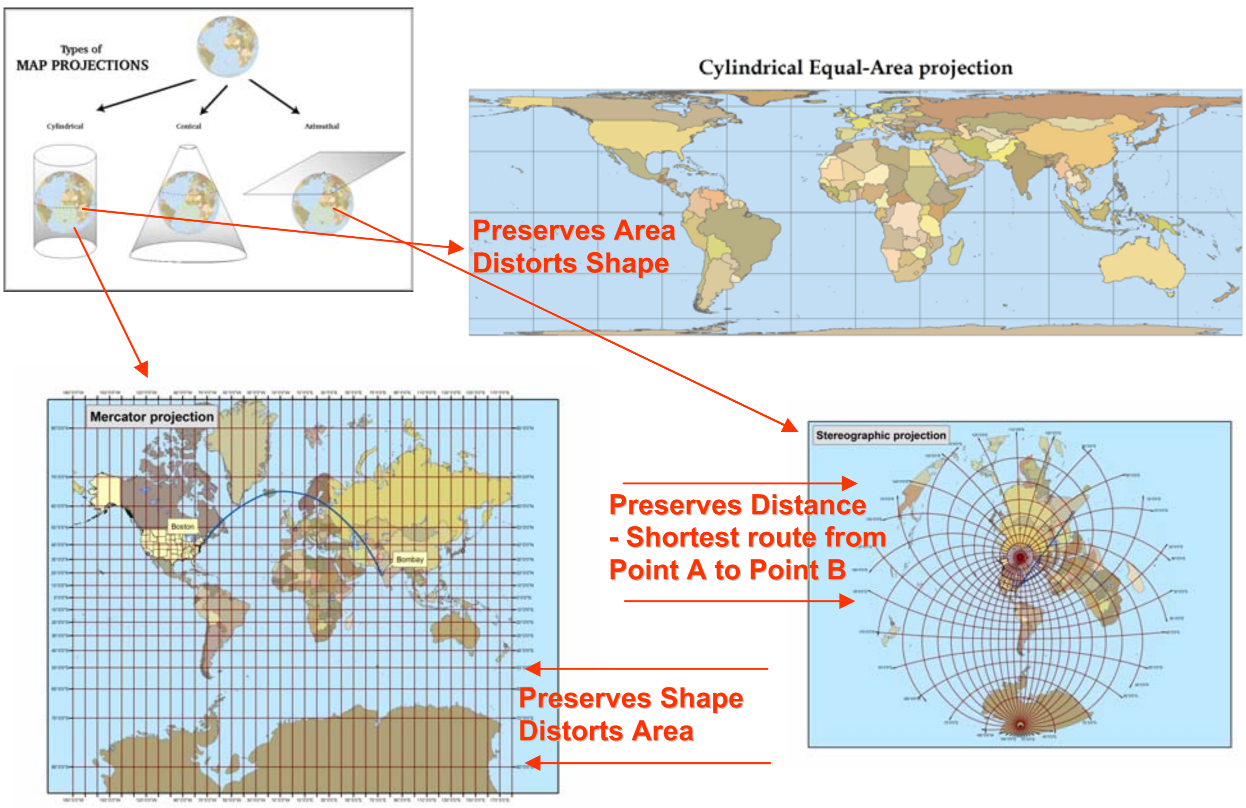


Figure 3: Various map projections **-** Mercator, Cylindrical Equal-Area, Stereographic - all exhibit some form of distortion

Map projections are applied to both maps and GIS data, to transform the 3d, spherical Earth surface into 2d. Projections are not absolutely accurate representations of the Earth. By necessity when flattening a rounded surface, some distortion occurs with direction, angle, shape and/or area. Map projections vary in how they distort/preserve these properties.

The Mercator projection preserves shape while greatly distorting area… Greenland is actually only 1/8 the size of South America. Area is preserved with the equal-area projections, though shape is distorted.

When a map is projected, the scale may vary across the map due to distortion; particularly when the map is portraying large areas such as the entire world or a continent. The scale parameter is entirely accurate only at the locations where the flat surface (cone, cylinder, plane) transects or touches the sphere. For cones and cylinders, the transect is a line called the **standard parallel**. For world maps, the equator is usually designated as the standard parallel. The further from the standard parallel(s) or azimuth, the greater the distortion for the map scale and other parameters (shape, direction, angle, area) that are not preserved with a particular parameter. The central meridian is along the prime meridian.

The type of projection we choose depends on the purpose of the map or analysis task.

There are numerous other map projection options and with a few clicks of the mouse in ArcGIS or other GIS software you could apply various options to basic world datasets that come by default with the software.

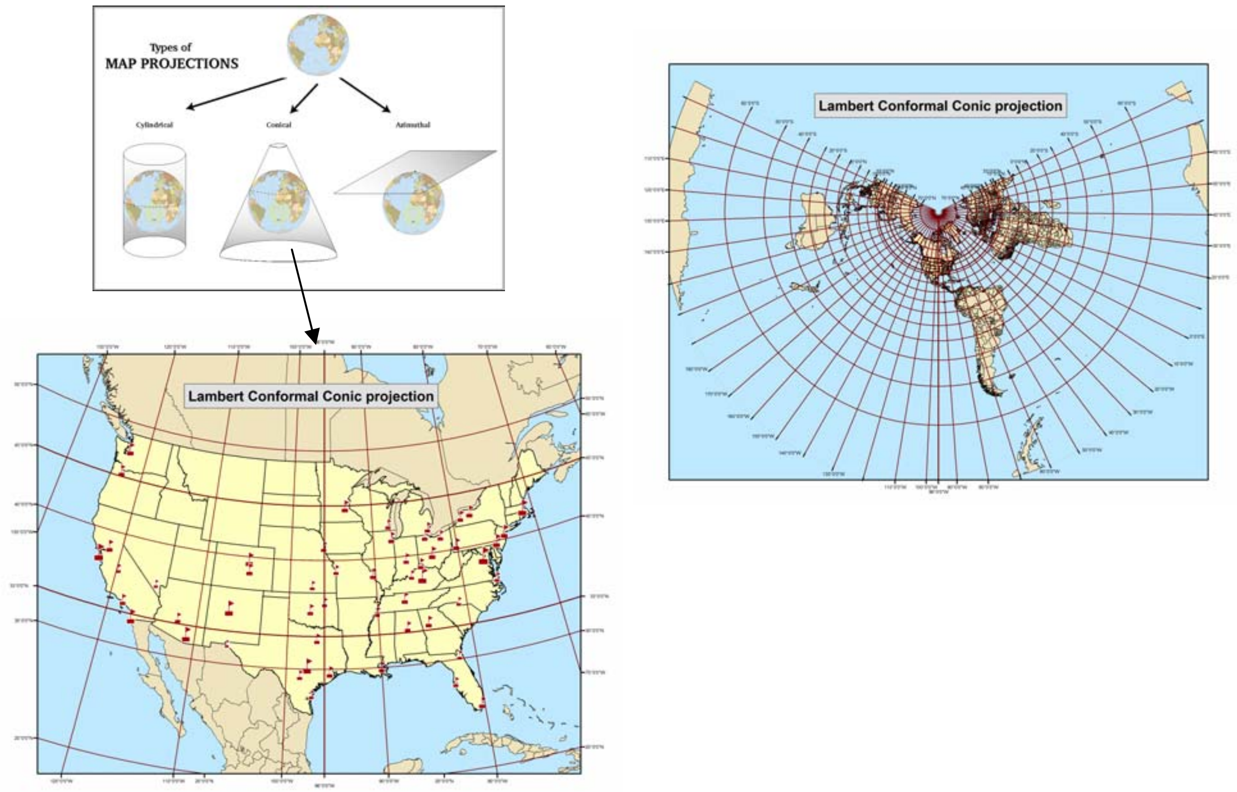


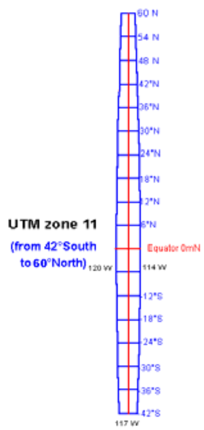
Figure 4: Lambert Conformal Conic projection, using two standard parallels, is suitable for maps of the United States**.**

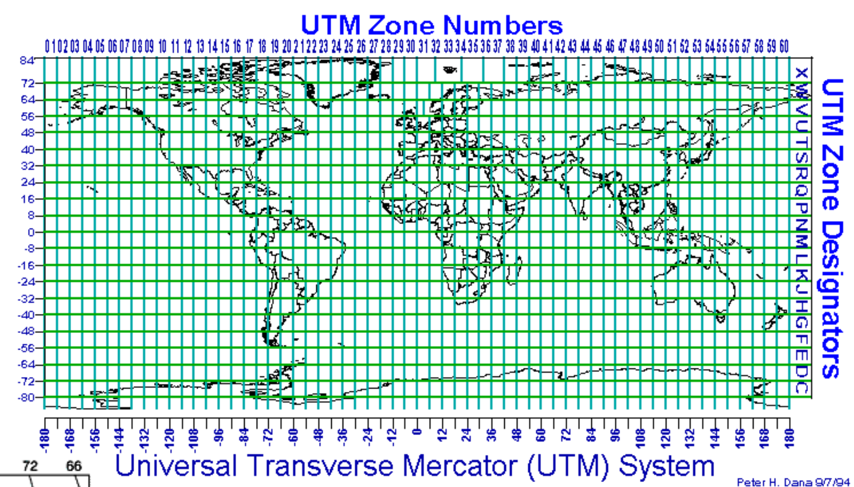
Lambert conformal conic projection is suited for smaller portions of the Earth, such as countries and is most suitable for countries or areas oriented east-west and in middle-north or south latitudes. As a conformal projection, it maintains angular conformity at the expense of preserving area. The parallels and meridians intersect at 90 degree angles.

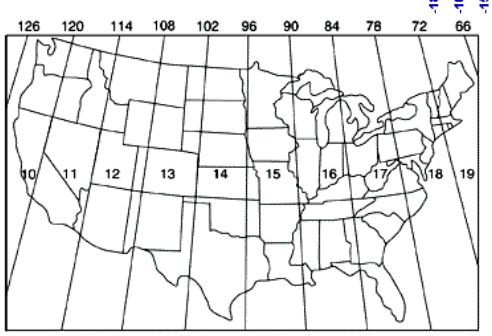
This projection moves the central meridian from the Prime Meridian, running through Britain, to 96 W longitude, running through the central U.S. This projection uses two standard parallels centered on the U.S., instead of the equator. The cone, instead of being tangent at one latitude, it 'cuts' through the Earth and intersects with the Earth's surface at two latitudes, giving us the two standard parallels. Area and shape are distorted away from the standard parallels. And by centering the standard parallels on the U.S., the distortion for the continental U.S. is minimized (though still present). However, this projection is totally unsuitable for say Australia!

Similar to the Lambert Conformal Conic, with two standard parallels, is the Albers Equal-Area conic projection. This is very suitable for a map of the continental U.S. It preserves the area at the expense of angles. Thus, preserving both angles and area is impossible, since they have an inverse relationship.

UTM is a system that divides the world up into 60 north-south zones, each 6 degrees of latitude in width. Each zone has its own central meridian and the origin is at the equator. UTM is based on the cylindrical, conformal, transverse Mercator projection. To eliminate negative coordinates, the coordinate system alters the coordinate values at the origin. The value given to the central meridian is the false easting, and the value assigned to the equator is the false northing. A false easting of 500,000 meters is applied. A north zone has a false northing of zero, while a south zone has a false northing of 10,000,000 meters. By centering the origin on the zone, we reduce distortion for particular areas. Zones 10 through 19 cover the United States.







Universe Transverse Mercator (UTM)

Lambert conformal conic projection forms the basis for the State Plane Coordinate System, used in the United States, for much smaller areas (for a state or portion of a state) that are oriented east-west such as Tennessee. Instead of the standard parallels running across the center of the United States, they would run across the center of Tennessee and thus provide minimal distortion for state and local mapping and GIS projects in Tennessee.

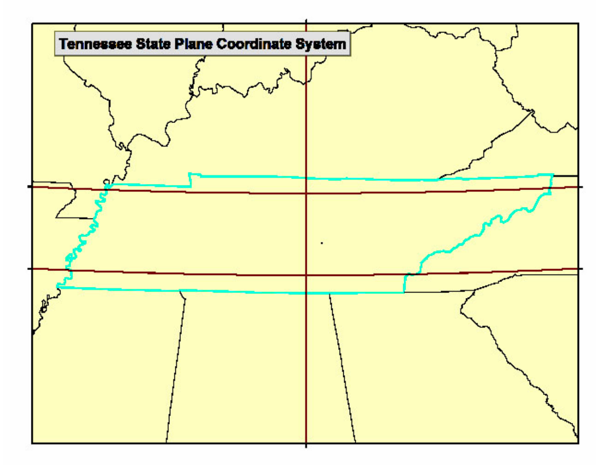


Figure 5: Tennessee State Plane Coordinate System centers the standard parallels and central meridian on the state, thus minimizing distortion.

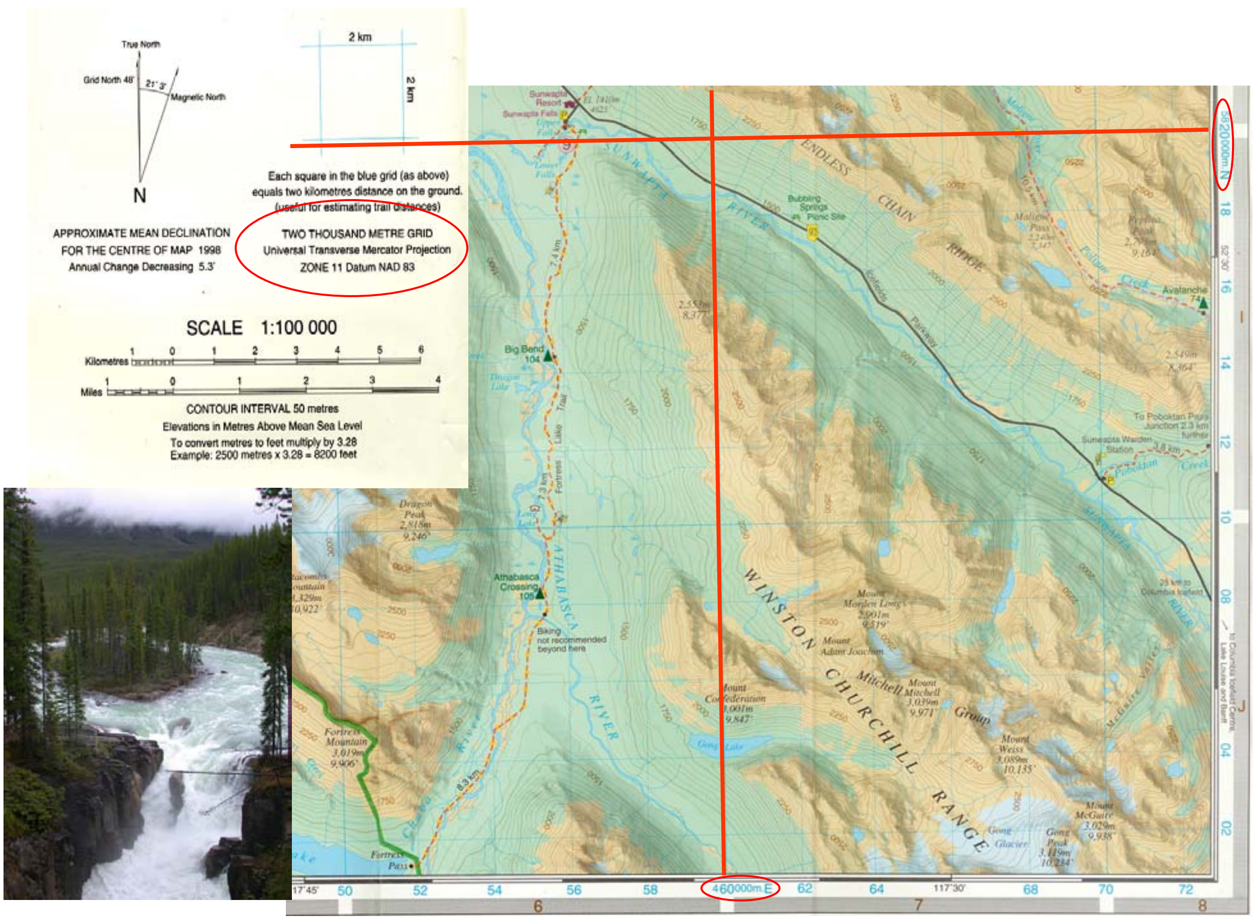


Figure 6: This map uses the UTM system. The coordinates on the map perimeter indicate that the map is 460,000 meters east and 5820000 meters north of the map origin (0 N, 120 W).

Back to this map of Jasper, the map uses UTM projection, zone 11. The map origin is at the equator (0 N) and 117 W. A false easting is applied, ‘moving’ the central meridian to 120 W so to eliminate negative numbers in the coordinates. So, 0,0 UTM coordinates are located at 0 N and 120 W. The coordinates on the map are indicated along the perimeter and by the map grid. This portion of the map is located 460,000 meters (or 460 kilometers) east or 120 W and 5,820,000 meters (or 5,820 kilometers) north of the Equator.

When combining different GIS data layers from disparate sources, it is important to make sure the layers are in the same projection and a related parameter, the datum, and the map units (feet, meters …) match. Though we won't go into details, common choice for datums in the U.S. is NAD83 as in this map. For projections, UTM or state plane coordinate systems are often chosen.

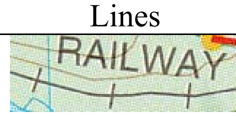
Bottom line is that when using data with mismatching map projections and datums, the GIS layers will not line up properly. In ArcCatalog, data/file management software, we can view this the projection and related information about the GIS datasets (metadata). As well, GIS software such as ArcGIS can easily convert GIS data from one projection to another.

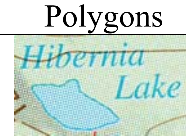
Symbolization

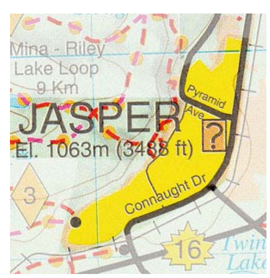
In a scale model we cannot possibly depict all the details and complexities of the real world. We must pick and choose which features to include in the GIS while omitting many lesser details. The features are represented on maps and in GIS as points, lines, polygons and surfaces. Furthermore, we might include a feature such as a road or river, yet cannot show all the intricate curves and detail. Instead we must generalize these features, straightening curves and reducing some features that in reality are small areas, to lines or points.

Maps and GIS use various symbols to represent features, and differentiate the features by varying the symbol colors, line widths, fill patterns, icons, etc. Discrete features (don't occur everywhere across the area) with abrupt boundaries (lake shores, roads, county boundaries…) are best represented in GIS as points, lines and polygons. This representation is the vector data model.

<table>
<tr><th>Points</th><th>Lines</th><th>Polygons</th></tr>
<tr>
<td>


</td>
<td>


</td>
<td>

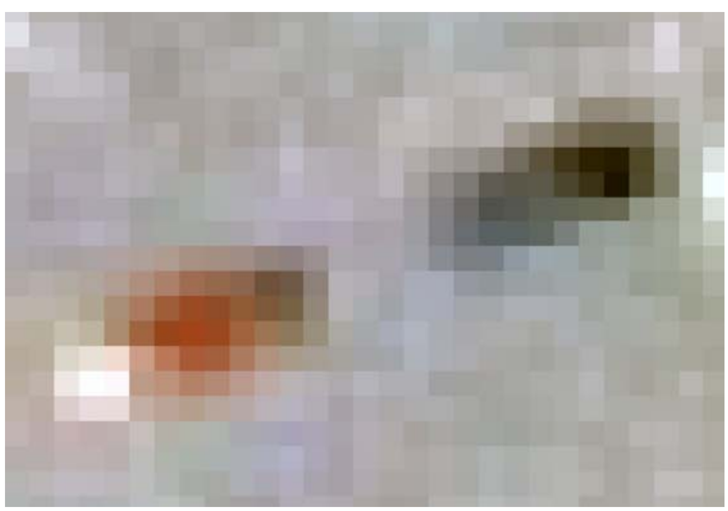


</td>
</tr>
<tr><td>Commercial Lodge
Backcountry camp</td><td>Railway
Road</td><td>Lake
River (at large scale)</td></tr>
</table>

Hostel Parking Lot Viewpoint	Hiking Trail River (at small scale)	Town (jurisdictions – city limits, counties, census tracts)

Alternatively, continuous features with gradual variations across the area are best suited for raster representation. Raster data uses a regularly spaced grid imposed on the 'map'. Each grid cell gives us the value for a particular feature or phenomena at that location. The phenomena, in an aerial photograph or satellite image, is the reflected light. It could also be surface elevation, temperature, among many other possibilities.

Raster:

Figure 7: Raster data is comprised of a continuous gridded surface, with each grid cell colored according to the value of the phenomena at that location. The phenomena here is simply the reflected light, but could instead by elevation, temperature, etc.



Thematic Mapping

While maps such as those above are useful for finding your way through the backcountry, for crime analysis and presenting crime statistics, thematic maps are more useful. They do not simply tell you, what's where? And the features are not necessarily mapped in proportion (true to map scale) to the features in reality.

If we are mapping major U.S. cities, we ought not to show them in direct proportion to the areas of each city. On this map, which shows just that, Jacksonville, Florida and Oklahoma City appear most prominent while Seattle, San Francisco and Boston barely exist.

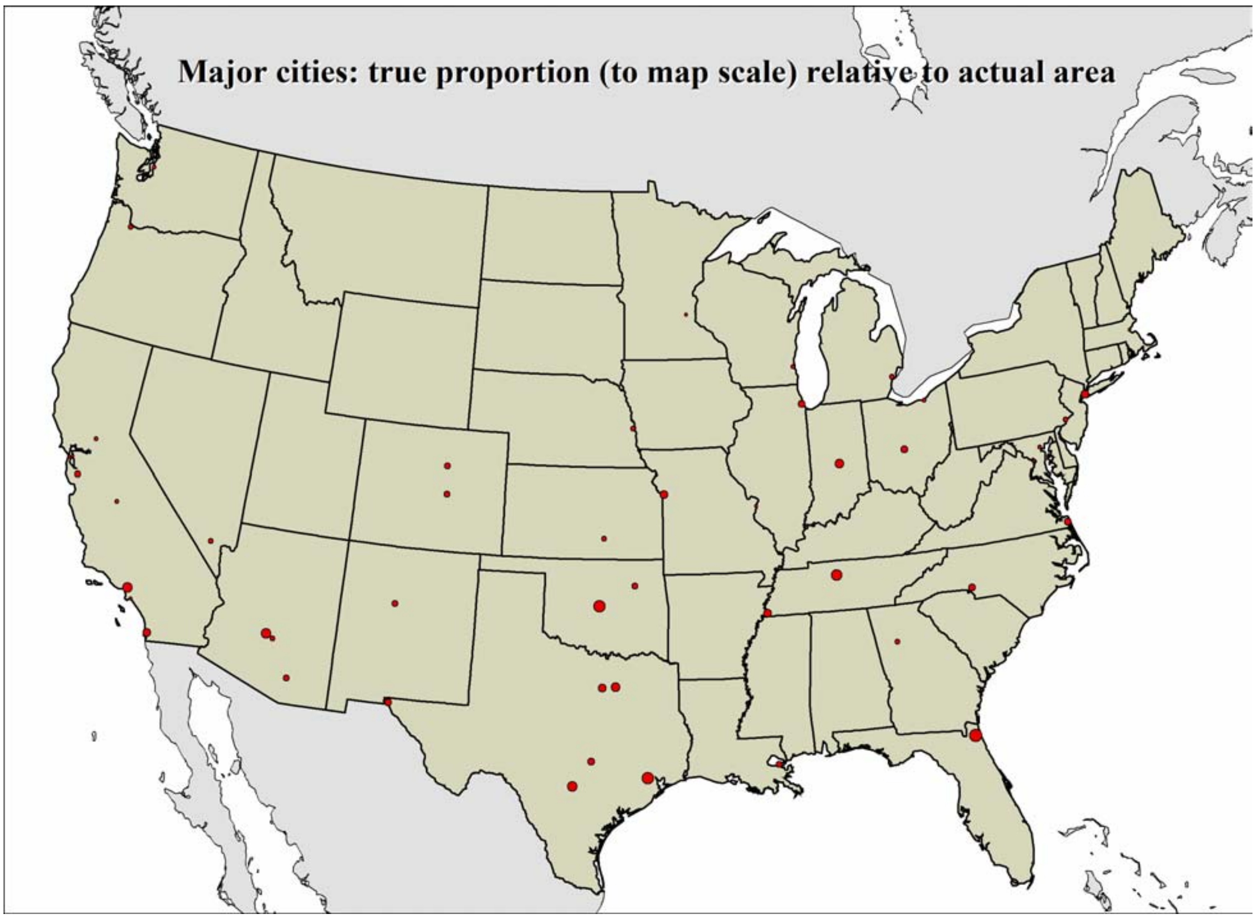


Figure 8: Major U.S. Cities in true proportion to their areas

Realistically, Boston, Seattle and San Francisco have large populations. As well, they are culturally and economically significant and should be depicted accordingly on the map, while Jacksonville and Oklahoma City can take a back seat.

Proportional symbols

Instead, thematic maps, using **proportional symbols**, adjust the point symbol sizes for the cities in accordance to some other variable such as population. This map uses proportional symbols for the metropolitan area population sizes. Taking into account the metro population is an even better indication of a city's size than the population within the city limits.

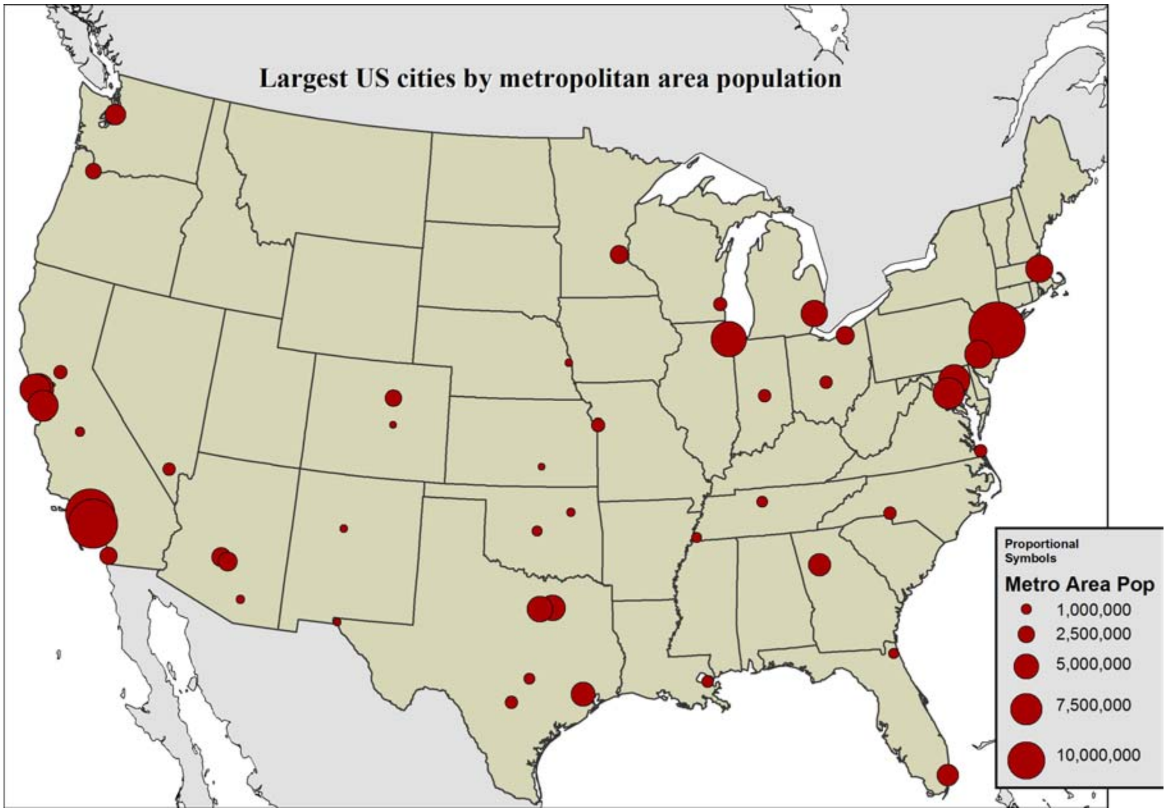


Figure 9: Major U.S. cities, proportional symbols according to metropolitan area populations.

Instead of representing major U.S. cities, proportional symbols can be used for police beats, census tracts, or other areas and instead of radio stations or education, can represent crime rates or other variables.

Data types (nominal, ordinal, interval, ratio)

Proportional symbols can only be used on quantitative attributes, such as population counts, crime rates, per capita income, etc. Other types of attribute data include nominal data, which are categories or named data. For U.S. States, Texas and Massachusetts are nominal attributes. We cannot say anything quantitative about these states based on the names alone. Nor can we say Texas is better than Massachusetts. The attributes are strictly names or categories. We could use unique or color categories for all the states, but the colors do not imply anything more.

Occasionally we may encounter ranked or ordinal data, for which we can use a graduated color scheme (light red/least favorite → dark red/most favorite).

Also, be careful when dealing with interval data, such as temperatures. Fahrenheit and Celsius temperature scales arbitrarily set freezing at 32 and 0 respectively. And we can no more say, 50 Fahrenheit is twice as much as 25 Fahrenheit.

Ratio data is any quantitative data such as population counts, number of crimes, crime rates, per capita income, etc. Proportional symbols and other thematic mapping is applicable only to such ratio (quantitative data).

Choropleth Maps

Choropleth maps instead use color shading schemes for areas (states, counties, census tracts …) to represent variables. Choropleth maps too, should use normalized variables than counts of some variable.

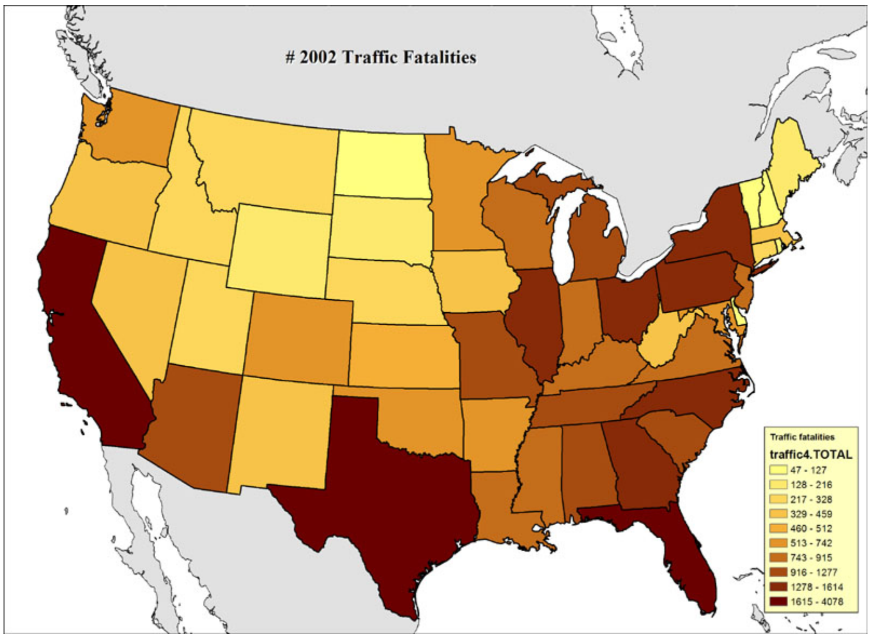


Figure 10: 2002 # of traffic fatalities

This map shows the number of traffic fatalities for each state in 2002. Mapping this way, of course California, Texas and New York have high numbers of traffic fatalities because they have large populations. Meanwhile, Montana and Wyoming roads appear quite safe.

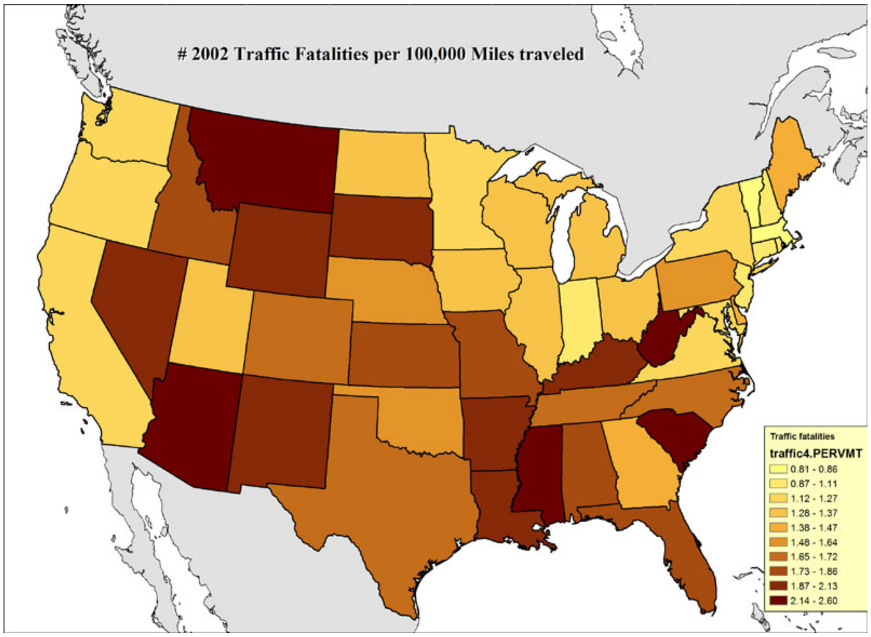


Figure 11: 2002 traffic fatalities per 100,000 miles traveled.

Data normalization

When mapping quantitative (ratio) data, such as counts (number of crimes, amount of sales, etc.), the data should be normalized. That is, average the counts (crimes, sales, etc.) by the population.

To get a better picture of the traffic fatality numbers, we should normalize the numbers. This map normalizes the fatalities by the miles traveled, while the map below normalizes by population. In these maps, Mississippi, South Carolina and Montana jump out, with high traffic fatality rates. Meanwhile California is quite safe and New York, along with Massachusetts has the lowest traffic fatalities rates in the United States. The non-normalized map can be misleading, reflecting more the states population than anything to do with traffic safety and fatalities.

When mapping crime statistics, be mindful of this. When mapping counts of crimes, you should normalize by the population of the area (police beat, census block, etc.) so you get a better picture of the crimes. Though, if the variable already comes as a mean or average, such as per capita income or crimes/100,000 population then those variables already are normalized. You can then use those averaged numbers straight as they are.

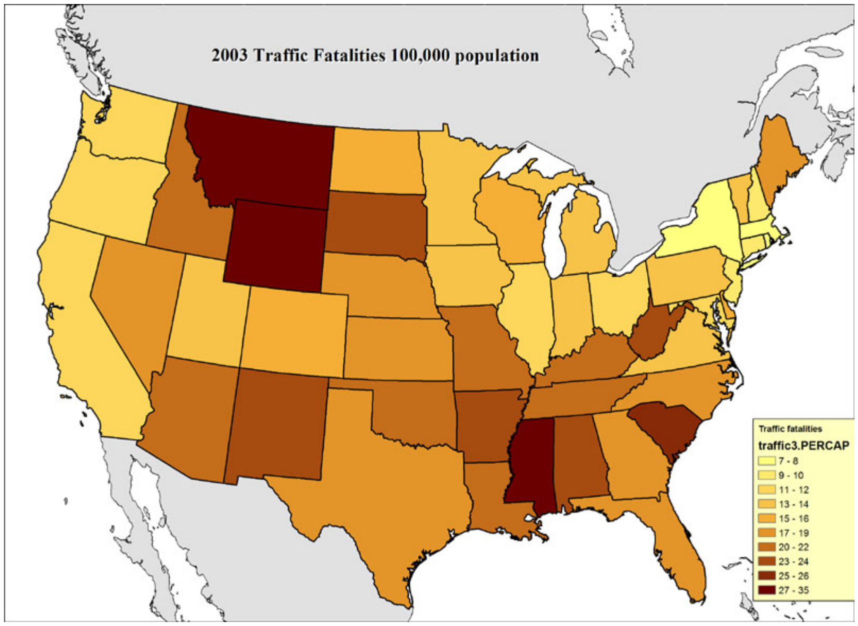


Figure 12: 2002 Traffic Fatalities per capita

GIS Analysis

Queries

Spatial queries allow you to search for features (crimes) within a particular area (police beat), crimes within 1000 feet of schools, etc. Where are the nearest subway stations to the hotel?

We can also combine spatial queries with tabular (attribute) queries. Which schools are elementary schools? Which crimes occurred within 1000 feet of these elementary schools?

Which subway stations serve the green subway line? Where are the nearest green-line subway stations?

Buffers

These zones, such as 1000 feet around elementary schools, are buffer zones. Querying features based on the buffer zones tells us which of the features are in proximity to elementary schools.

Local zoning ordinances often have specific buffer requirements for locating new bars, adult establishments and gambling establishments away from schools, libraries, residential areas, churches & other religious establishments, and day cares.. We can use GIS to identify all such places and establish a 1000 ft buffer (or whichever distance specified in the ordinances) to find locations that meet and do not meet this criteria.

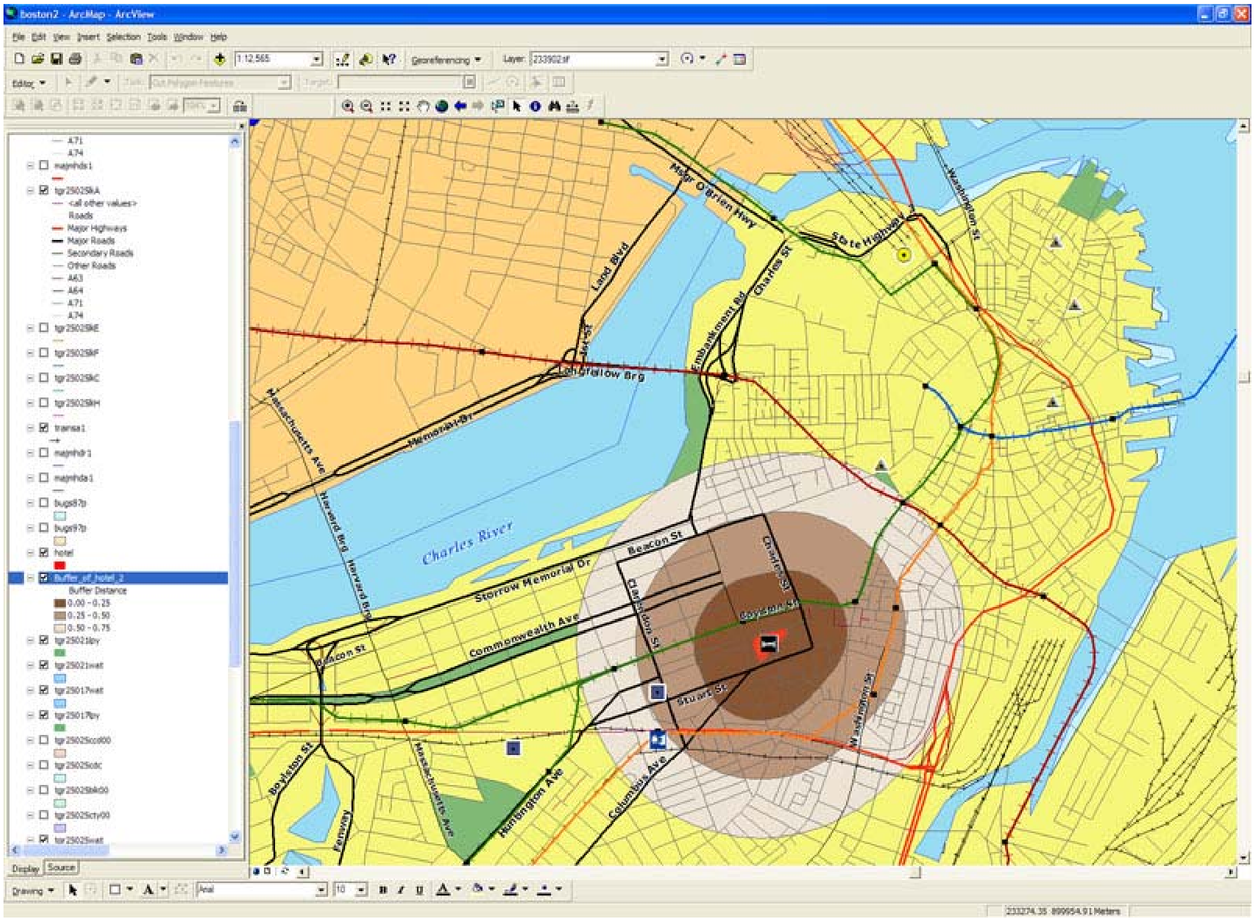
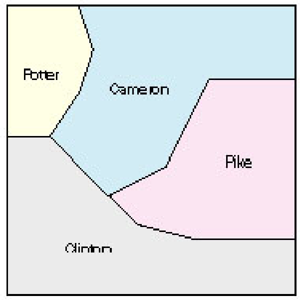
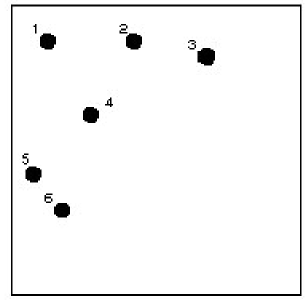


Figure 13: Buffer zone around the Boston Park Plaza Hotel. We can select subway stations, restaurants or other places that fall within this proximity/buffer zone.

Spatial Overlay

- Point-in-polygon
- Line-on-polygon
- Polygon overlay

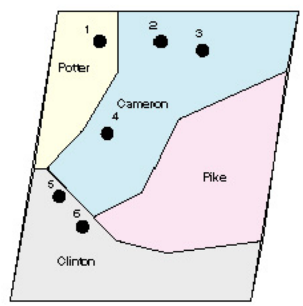


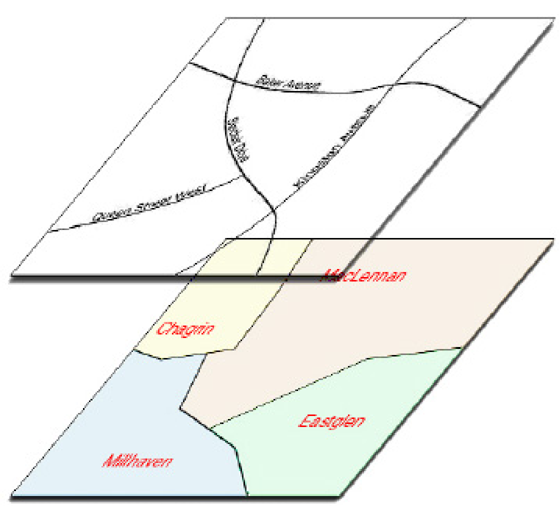


Polygon Layer indicates the four different counties being looked at for the possible building of a reservoir. The reservoir will feed the four county radius with fresh drinking water.

Point Layer indicates the documented points of contaminated ground water. The contaminations were found in local streams over the last ten years.

The final point in polygon layer shows what counties have had documented stream contamination in the last ten years. Hopefully this map will indicate that building a reservoir in Eastern Pike County would be the safest and best location.





Like overhead transparencies, these layers are placed on top of eachother, allowing you to see both school districts and roads together on one layer.

Spatial patterns (Clustered/Random/Regular)

The prior example deals with a small number of spatial objects. However, many features occur in far greater numbers (trees in a forest, crimes in a city, etc.), but are not necessarily uniformly located within these areas. Each set of objects exhibits a particular **spacing** or set of **arrangements**. These spacings seem to have underlying controls or processes that dictate their placement. Some do exhibit a regular, uniform spacing pattern, such as corn plants in a farmer's field.

Trees in a forest, for example, display a scattered or **random** pattern, with no apparent underlying design. Alternatively, some objects tend to cluster (geese in/around a pond, crime near illegal drug activity…). Clustered objects represent a high density of features, whereas other objects demonstrate a more dispersed pattern.

Spatial diffusion/displacement

Another important spatial process is diffusion…from the diffusion of infectious diseases such as West Nile Virus. **Spatial diffusion** patterns can be divided into four major categories: *expansion diffusion* characterizes that the spreading phenomena, such as wildfire, has a source and diffuses into new areas. *Relocation diffusion* describes spreading phenomena that moves into new areas, leaving behind the original/source locations, such as that occurs with migration. *Contagious diffusion* occurs where the spread requires direct contact among individuals, such as with infectious diseases. Finally, *hierarchical diffusion* describes spread through sequences of classes or places, such as a new fad that spreads from major urban centers to smaller towns.

GIS in Law Enforcement

http://www.ojp.usdoj.gov/nij/maps/

Provides listing of some local law enforcement crime mapping on the web.

Geographic Information Systems (GIS), along with related spatial and information technology, can enable law enforcement officers to get the right information (suspect's criminal history record across this and other jurisdictions? Is he/she on a FBI watch list? etc) at the right time (immediately) and at the right place (in the field, at the crime scene or in pulling over a driver). It is also highly beneficial to employ GIS for crime analysis and investigations, being able to integrate a wide range of spatial information alongside non-spatial information and analyze relationships among variables and factors. As well, in knowing where the crime hot spots are, police management can better allocate a department's resources to deal with problems. And related, GIS is integral for emergency response (9-1-1) systems for dispatching officers and emergency personnel. There are certainly many more applications of GIS for law enforcement and criminal justice, though these are currently among the most commonly deployed uses of GIS.

Benefits of GIS for crime mapping:

The ability to overlay disparate data sets makes it an excellent tool for identifying factors related to multidimensional, multi-faceted crime problems.

Three primary goals:

- To further understanding of the nature and extent of criminal and social problems in a community, particularly the relationship between criminal activity and possible contributing factors. (**crime analysis**)
- To improve allocation of resources to combat these problems. (**resource allocation**)
- Inform the public regarding crime trends, statistics and information.

Citizen awareness and information

Some jurisdictions employ GIS to provide their citizens with these services:

- Citizens can find their police district, police service area and officers (including their e-mail and phone #) serving their neighborhood, by entering their street address or browsing the information.
- Crime statistics by district and PSA can be provided
- Crime statistics/mapping
- Sex offender registry (locate sex offenders near an address)

Getting Started with GIS

GIS courses

- Colleges/universities
  - http://www.gis.psu.edu/education/
- Vendor training
  - On-site
  - E-learning (http://campus.esri.com/)
- Workshops
  - NLECTC (National Law Enforcement and Corrections Technology Center)

Recommended GIS Reading

- How to Lie with Maps (Mark Monmonier)
- Principles of Geographic Information Systems (Peter A. Burrough)
- MAPS online references (http://www.ojp.usdoj.gov/nij/maps/)

Get assistance in pursuing GIS projects

- Other GIS/IT projects in your jurisdiction
- Neighboring jurisdictions pursuing crime mapping
- Assistance from state governments
  - Massachusetts GIS Clearinghouse (http://www.state.ma.us/MGIS/)
  - Justice IT integration programs
  - Possible funding?
- Federal assistance
  - NIJ-MAPS
  - Geodata.gov initiative
  - Grants from Homeland Security, Justice & other departments
- GIS consultants/vendors

Links

NIJ/MAPS program: http://www.ojp.usdoj.gov/nij/maps/

Interactive Crime Mapping

Lincoln, Nebraska crime mapping: http://www.ci.lincoln.ne.us/city/police/
Memphis crime maps: http://www.memphispolice.org/
Oakland crime maps: http://www.oaklandnet.com/cw/MapMain.jsp
Salt Lake City crime maps: http://www.slcgov.com/police/online.htm

Macromedia Flash Crime Maps

Boulder, Colorado – crime maps: http://www.ci.boulder.co.us/police/crime/crime_map_flash.htm

Static Crime Maps
Tempe, Arizona – crime maps: http://www.tempe.gov/cau/crime_analysis_maps.htm

Sex Offender Registries
Miami-Dade crime maps, sex offender registry and other maps (health code violations, transportation …):
http://gisims2.co.miami-dade.fl.us/MyNeighborhood/home.asp

Riverside, California sex offender registry: http://www.co.riverside.ca.us/sheriff/crime/meganlaw.htm

GIS Education
Penn State GIS Distance Learning: http://www.gis.psu.edu/education/
ESRI Virtual Campus: http://campus.esri.com/
NLECTC: http://www.nlectc.org/cmap/justnet.html

Contact Info

Katie Filbert
National Institute of Justice, MAPS
810 7th St NW
Washington D.C. 20531

Phone: (202) 305-7530
E-mail: filbertk@ojp.usdoj.gov

http://www.ojp.usdoj.gov/nij/maps/